Cryo-EM structure of SARS-CoV spike glycoprotein in post-fusion state

Xiaoyi Fan[1,2,4], Duanfang Cao[1,4,*], Lingfei Kong[1] & Xinzheng Zhang[1,2,3,*]

Affiliations:

[1]National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, P.R. China.

[2]University of Chinese Academy of Sciences, Beijing 100049, P.R. China.

[3]Center for Biological Imaging, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, P.R. China.

[4]These authors contributed equally to this work.

[*]Correspondence to : caodf@ibp.ac.cn (D.C.); xzzhang@ibp.ac.cn (X.Z.)

Abstract

Global emergencies caused by the zoonotic severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV) and the newly discovered 2019 novel coronavirus (2019-nCoV) have posed a serious threat to human health. The spike (S) glycoprotein, a homotrimer located on the surface of the viral envelope, is the key antigen for therapeutic development. The S2 subunit from the S glycoprotein, which is highly conserved among coronaviruses, contributes to viral entry by mediating the host-viral membrane fusion. However, the structural information of the post-fusion S2 machinery from these highly pathogenic human-infecting coronaviruses is still lacking. Here, we report the structure of SARS-CoV S glycoprotein in the post-fusion state by single particle cryo-electron microscopy, revealing a more rotated HR1-HR2 six-helix bundle and a tightly bound linker region upstream of the HR2 motif that plays an important role in membrane fusion. Comparison with the structure of pre-fusion SARS-CoV S glycoprotein shows dramatic structural rearrangements and conformational changes, resembling that of the Mouse hepatitis virus (MHV) and other class I viral fusion proteins. By analyzing the structural features, describing the glycan shield and mapping the antibody and inhibitor targets on the surface of the post-fusion S glycoprotein, we provide structural basis of potential therapeutic targets within the highly conserved S2 subunit, which may help the development of effective vaccines and therapies against a wide range of SARS-like coronaviruses.

Keywords: SARS-CoV; Spike; post-fusion; cryo-EM structure; glycan shell

Introduction

Coronaviruses are a family of enveloped, positive-sense, single-stranded RNA viruses that usually cause mild respiratory tract infections in humans[1]. However, the outbreaks of severe acute respiratory syndrome coronavirus (SARS-CoV) infection and Middle-East respiratory syndrome coronavirus (MERS-CoV) infection in the past two decades caused serious epidemics and brought the two highly pathogenic human-infecting coronaviruses to the forefront[2]. SARS-CoV infection that emerged from China to the global world in 2002 resulted in a lethality rate of 10%, and MERS-CoV that was first discovered in Saudi Arabia is responsible for approximate 2,500 infections and 850 deaths since 2012[3-5]. Then, the outbreak of epidemic pneumonia in Wuhan of China since December 2019 caused by a newly discovered 2019 novel coronavirus (2019-nCoV) have led to about 77,000 infections and 2,400 deaths until February 23th 2020[6-8]. The 2019-nCoV infection has spread within China and emerged globally to other countries, but effective vaccines and therapeutics are still lacking. Moreover, SARS-CoV, MERS-CoV and 2019-nCoV are all zoonotic coronaviruses that possibly originate from bats and transmit to humans through different intermediate hosts[2,9,10]. Considering the possibility of continuous contact between humans and the intermediate hosts, and the high mutation rate of RNA viruses, additional emergencies caused by theses coronaviruses or other undiscovered coronaviruses have become a serious potential threat to global public health[9]. Thus treatments toward a wide range of coronaviruses are urgently needed.

The key determinant for efficient animal-to-human and human-to-human transmissions of coronavirus is the spike (S) glycoprotein located on the virus surface, which is responsible for viral entry[2,11]. The coronavirus S glycoprotein is a class I viral fusion protein and exists as a homotrimer with multiple glycosylate modifications. The S glycoprotein is composed of an N-terminal S1 subunit mediating host cell receptor recognition and a C-terminal S2 subunit responsible for viral and host cell membrane fusion process[2]. Cleavage at the boundary of S1 and S2 subunits and interaction with the host receptors induce the dissociation of the S1 subunits from the spike[12,13]. The remaining S2 trimer, which is highly conserved among

coronaviruses, undergoes a series of conformational changes to trigger membrane fusion between the target cell membrane and the viral envelop[11]. Because of the critical role of S glycoprotein in viral transmission, S glycoprotein is the major target of neutralizing antibody and the key factor for drug development[2,14]. Antibodies and inhibitors against both the S1 and S2 subunits have been reported to inhibit viral entry[14-20]. According to the high sequence identity of the S2 subunit among coronaviruses, the S2-specific therapies are believed more suitable for treatment of a broader range of coronavirus infections[18,21].

Previously, we and others have determined the pre-fusion structures of the S glycoproteins of SARS-CoV and MERS-CoV, and revealed the dynamic receptor binding domain (RBD) with a standing conformation responsible for receptor binding, which is unique to the two highly pathogenic human-infecting coronaviruses[15,22-24]. Similar phenomenon has been conformed in 2019-nCoV[25]. The structure of the SARS-CoV S glycoprotein in complex with ACE2 further demonstrated the interaction between the standing RBD and the receptor, which may facilitate the release of the S1 subunits[26,27]. Structural information of the S2 subunit from SARS-CoV or MERS-CoV in hemi-fusion or post-fusion state is relatively limited, except for the heptad repeat (HR) motifs, which fold into an HR1-HR2 six-helix bundle as shown by their crystal structures[28-31]. Meanwhile, a research group determined both the pre-fusion and post-fusion structures of mouse hepatitis virus (MHV) S glycoprotein, and described the structural rearrangements of MHV S2 subunit during membrane fusion[32,33]. However, detailed structural information of the SARS-CoV, MERS-CoV as well as 2019-nCoV S2 fusion machineries are still needed for further understanding the membrane fusion process and broad-spectrum therapeutics development against these human-infecting coronaviruses with serious pathogenicity.

Here we report the cryo-electron microscopy (cryo-EM) structure of SARS-CoV S2 trimer in post-fusion state, and present the N-linked glycan shield masking the surface of the S2 machinery. Comparison with the structure of SARS-CoV S glycoprotein in pre-fusion state reveals dramatic conformational changes of SARS-CoV S2

machinery during membrane fusion, resembling that of MHV and other class I fusion proteins. In the structure of the post-fusion SARS-CoV S2 machinery, we found that the linker regions upstream of the HR2 motifs that critical for HR1-HR2 six-helix bundle formation bind more tightly along the central stem region than that from MHV, and the HR1-HR2 six-helix bundle are more twisted leading to a less buried area between the HR1 and HR2 motifs. By mapping the antibody and inhibitor targets within the conserved S2 subunit, we provide the structural basis for development of vaccines and therapies which may suitable for treatment of a broad range of SARS-like coronaviruses.

Methods

Protein expression and purification

The gene of SARS-CoV spike protein ectodomain (UniProt accession number P59594, residues 14–1,193) was optimized for synthesis and inserted into the baculovirus transfer vector pFastbac1 (Invitrogen) with an N-terminal gp67 signal peptide, a C-terminal thrombin cleavage site followed by a T4 fibritin trimerization domain and a His_6-tag. Transfection and virus amplification were performed in Sf9 cells, and Hi5 cells were used to express the recombinant proteins using the Bac-to-Bac system. The recombinant proteins were purified as previously described[22]. Briefly, the proteins were harvested from cell culture medium, and purified sequentially on Ni-NTA column and Superdex 200 column (GE Healthcare).

Human ACE2 extracellular domain (UniProt accession number Q9BYF1, residues 19–615) with an N-terminal gp67 signal peptide and a C-terminal His_6-tag was optimized for synthesis and subcloned into baculovirus transfer vector pFastbac1 (Invitrogen). Transfection and virus amplification were performed in Sf9 cells, and Hi5 cells were used to express the recombinant proteins using the Bac-to-Bac system. The proteins were harvested from cell culture medium, and purified sequentially on Ni-NTA column. Fractions containing ACE2 were collected and applied directly to QHP column (GE Healthcare), and then eluted with a 0.15–1 M NaCl gradient in 20 mM Tris buffer (pH 8.0). Fractions containing ACE2 were then purified using a

Superdex 200 column with a buffer containing 20 mM Tris–HCl (pH8.0) and 150 mM NaCl.

Induction of S2 subunit in post-fusion state

The purified S proteins were incubated with trypsin (Sigma, trypsin with S protein mass ratio of 1:100) and thrombin (Sigma, 3 units per mg S protein) overnight at room temperature to cleave at the S1/S2 boundary and remove the C-terminal trimerization domain and the His_6-tag. The size exclusion chromatography was conducted to purify the cleaved product with a Superdex 200 column. The S1 and S2 subunits still formed stable complexes according to the results from Gel-filtration chromatogram and SDS-PAGE analysis. Then the concentrated sample was mixed with 1/10 volume of 1 M sodium citrate at pH 5.5, $CaCl_2$ with a final concentration of 2 mM, and human receptor ACE2 (S protein with ACE2 at molar ratio of 1:20), followed by incubation at room temperature for 16 hours. The mixture was purified by gel-filtration chromatography using a Superdex 200 column and analyzed by SDS-PAGE.

Cryo-electron microscopy data collection and processing

Purified protein sample with a concentration of 0.1 mg/ml was placed on a glow-discharged holey carbon grid (Quantifol, Au 400 mesh, R 0.6/1.0). After 4 s blotting with filter paper, the grid was flash plunged in liquid ethane using an automatic plunge device (Leica EM GP2) with 10℃ temperature and 75% humidity. Cryo-EM single particle data collection was performed through our newly developed beam-image shift data collection method[34] using a 200 kV FEI Tecnai Arctica microscope equipped with K2 camera (Gatan). Images were recorded at a defocus range of -1.8 μm to -2.2 μm with a pixel size of 1.0 Å. The exposure time was 5.12 s, with a total exposure dose of ~50 electrons per $Å^2$ over 32 frames.

Each image stack was subjected to motion correction using the software MotionCor2[35] before further data processing and resulted in an averaged micrograph. The parameter of the contrast transfer function on each micrograph was determined

by the program CTFFIND4[36]. All followed up processing steps were executed in RELION 2.1[37] (Figure S2). A subset of the protein particles was manually picked and processed with reference-free 2D classification. Three representative 2D class averaged images were selected as references for particle automatic picking of the whole dataset, which resulted in a total of 343,042 particles. Then the particles were processed by reference-free 2D classification. After two rounds of 2D classification, 69,792 particles were selected for further 3D classification, which were classified into five classes using the MHV S2 post-fusion trimer density map[33] as an initial model. 40,600 particles from class 4 were subjected to further 3D auto-refinement with C3 symmetry, yielding a 3.9 Å density map estimated based on the gold-standard Fourier shell correlation with 0.143 criterion. The local resolution of the final density map was calculated using ResMap[38].

Model building and refinement

For model building of the post-fusion SARS-CoV S2 trimer, the structure of SARS-CoV S2 subunit was predicted by Phyre2[39] using the post-fusion MHV S2 trimer (PDB : 6B3O)[33] as template. Then, the predict structure was fitted into the 3.9 Å resolution map using UCSF Chimera[40] and manually rebuilt in Coot[41]. The model was further improved by cycles of real space refinement in Phenix[42] with geometry and secondary structure restraints, and subsequent manual corrections by Coot were carried out iteratively. The refinement statistics of the model generated in Phenix is summarized in Table 1. Figures representing the structural features were prepared with UCSF Chimera[40] and PyMOL (http://pymol.org).

Results

Architecture of SARS-CoV S2 fusion machinery in post-fusion state

We produced the post-fusion SARS-CoV S2 subunit by stimulating the pre-fusion S ectodomain with trypsin digestion, receptor binding and low pH treatment, and determined the cryo-EM structure of the induced post-fusion SARS-CoV S2 trimer at a resolution of 3.9 Å (Supplementary Fig. 1 and 2). The final structure includes all the

domains and regions of the S2 subunit (residues 688-1178), except for the fusion peptide (FP) and its downstream connecting region (CR) motifs (residues 754-900). Similar to the structure of post-fusion MHV S2 subunit generated by directly expressing only the S2 subunit ectodomain[33], overall structure of the SARS-CoV post-fusion S2 trimer adopts a baseball bat shaped appearance with dimensions of approximate 50 Å by 20 Å by 185 Å (Fig. 1). The structure of post-fusion SARS-CoV S2 fusion machinery comprises a long central helical bundle surrounded by short helixes and β-sheets at the membrane distal end (Fig. 1). The HR1 and central helix (CH) motifs fold into a continuous 31-turn α-helix. Three HR1-CH long helixes intertwine with each other and form the stem region of the post-fusion S2 trimer with an orientation outward from the membrane side. The HR2 motif and its upstream linker region bind into the groove at the interface between the two HR1-CH helixes from the neighboring protomers in an antiparallel manner, reminiscent of the six-helix bundle described in the crystal structures of HR1 and HR2 motifs[28-31]. The upstream helix (UH), β-hairpin (BH) motif and subdomain 3 (SD3) are located surrounding the C-terminal of the central HR1-CH helixes, which only partially cover up the top of the central helical region.

Transition from pre-fusion to post-fusion state triggers dramatic structural rearrangements and conformational changes of the SARS-CoV S glycoprotein (Fig. 2), resembling that of MHV S, Influenza virus HA, paramyxovirus F and other known class I viral fusion machineries[33,43,44]. The largest structural rearrangement lies on the HR1-CH helical region (Fig. 2c). In the pre-fusion state, the HR1 and CH regions are constituted by separate short helixes, and the HR1 helixes and the CH helix are connected by a U-turn loop, which makes the two motifs pose to antiparallel orientation. After membrane fusion, the HR1 motif turns over and forms a continuous stem helix together with the CH motif, pointing to the orientation toward the target membrane, which makes the post-fusion SARS-CoV S2 fusion machinery extends approximate 80 Å longer compared to the pre-fusion S2 trimer. Note that all the flexible connecting loops from the HR1 and CH motifs in the pre-fusion state rearranged into the long single stem helix in the post-fusion state. Such an onward

helical extension contributes to the ejection of the fusion peptide. The UH, BH and SD3 regions retain their tertiary structures in the pre-fusion state at the viral membrane proximal side, but the UH and SD3 regions rotate slightly relative to the CH helix, making the quaternary structure a little bit compact (Fig. 2d and Supplementary Fig. 3a-c). As described in the post-fusion MHV S2 structure[33], four highly conserved disulfide bonds (C720-C742, C725-C731, C1014-C1025 and C1064-C1108) are identified within the post-fusion SARS-CoV S2 trimer, which also exist in the pre-fusion state, are located exactly within these relatively stable regions (Supplementary Fig. 3). These disulfide bonds, especially the linkage between domains, help to maintain the structure of the S2 subunit during the dramatic transition triggering membrane fusion.

The upstream linker region of the HR2 motif critical for membrane fusion

The upstream linker region of the HR2 motif resolved in our post-fusion structure is only partially visible in the structure of pre-fusion SARS-CoV and MERS-CoV S glycoproteins, but could be completely observed in the structure of pre-fusion human coronavirus NL63 (HCoV-NL63) S trimer[15,45]. In the pre-fusion state, a β-strand from the pre-fusion linker region forms a β-sheet structure with the SD3 domain, and thus the upstream linker region is well folded and retained adjacent to the SD3 domain at the viral membrane proximal side (Fig. 3c-e). By contrast, in the post-fusion state, the linker region no longer folds into the β-sheet with the SD3 domain, but is rearranged into an extended structure mainly assembled by loops and binds along the axis of the stem helixes from adjacent protomers in the orientation toward the target membrane (Fig. 3). The N-terminal of the linker region is stabled by interaction with the hydrophobic surface of the core β-sheet and the subsequent hydrophobic groove formed by UH helix and CH helix (Fig. 3b). The C-terminal region binds to a positive charged cavity through its acidic amino acids enriched region (residues 1128-1133) (Fig. 3b). This observation indicates that the linker region upstream of the HR2 motif also undergoes large structural rearrangements during membrane fusion, and right refolding of this linker region is important for the formation of the central HR1-HR2

six-helix bundle, which brings the viral membrane close to the host membrane at the late stage in the fusion transition.

However, conservation analysis shows that while all the first three disulfide bonds are completely conserved between SARS-CoV, 2019-nCoV, MERS-CoV and MHV, the fourth disulfide linkages (C1064-C1108) between the upstream linker region of HR2 motif and the SD3 domain is not exist in the MERS-CoV S glycoprotein, as the residue corresponding to C1108 in the linker region is replaced by a glutamine (Supplementary Fig. 3d). Considering that this disulfide linkage is the last contact between the β-sheet region from SD3 and the N-terminal end of the linker region in the post-fusion structure, this linkage as well as the rotation of the SD3 domain contribute to pose the linker region in the right direction along the axis of the stem helix toward the target membrane, and thus benefits subsequent binding of the linker region and HR2 motif into the groove formed by the HR1-CH helixes (Fig. 3f). Moreover, S glycoprotein has to finish its fusion transition as soon as the dissociation of the S1 subunits, or the exposed fusion motifs may be attacked by neutralizing antibodies from host immune system. Thus this additional disulfide bond in SARS-CoV and 2019-nCoV may contribute to higher transmission efficiency among humans.

Comparison with S2 fusion machineries from other coronaviruses

Consistent with the high sequence identity (40%) between SARS-CoV and MHV S2 subunits, the post-fusion structure of SARS-CoV and MHV S2 trimers adopt similar architecture with an overall root mean square deviation (RMSD) of 2.43 Å (Fig. 4). Structure superposition analysis showed that the SARS-CoV HR2 helixes twist approximately 4 Å around the central HR1 helixes compared to that of MHV, while the HR1-CH helixes rotate slightly (Fig. 4a-c). This rotation forces the SARS-CoV HR2 motifs (residues 1145-1178) less buried by the groove assembled from the HR1 helixes, with a buried surface area of 5162 $Å^2$, compared to 5780 $Å^2$ between the MHV HR2 motifs (residues 1215-1248) and the grooves. By contrast, the SARS-CoV linker region (residues 1105-1143) upstream of the HR2 motif binds tightly along the

central stem helix, while the MHV linker region (residues 1174-1212) shifts a little bit outward from the binding groove (Fig. 4c-e), which makes the accessible surface area of the MHV linker regions less buried for about 345 $Å^2$. These structural differences suggest that the SARS-CoV linker region upstream of the HR2 motif may contribute more to the six-helix bundle formation than that in MHV, which makes the linker region upstream of the HR2 motif an attractive site for drug designation and antibody development to defense SARS-CoV infection.

Interestingly, as previously reported, superposition of the crystal structures of the HR1-HR2 six-helix bundles from MHV and SARS-CoV shows that the SARS-CoV HR2 helix in the crystal structure is rotated about 1.5-2.5 Å around the stem helixes compared to that of MHV[29], which is slightly less twisted than the HR2 motif in our S2 trimer structure (Supplementary Fig. 4). Lacking restriction and stabilization of other regions beside the HR1 and HR2 motifs in the crystal structure may account for these differences. The three CH helical regions forming continuous α-helixes with the HR1 motifs are more intertwined with each other in our post-fusion S2 machinery compared to the MHV CH helixes (Fig. 4a and 4c), and thus initiating the slight rotation of the refolded HR1 helixes (Fig. 4a-c) and further leading to twist of the HR2 motifs spontaneously. However, we compared the crystal structures of HR1-HR2 helixes from MERS-CoV and MHV, and found that their structures adopt similar architectures without any rotation (Supplementary Fig. 4c). Hence, these observations indicate that the rotation of the HR2 motif may be a unique structural feature for SARS-CoV and SARS-CoV related coronaviruses, for example, the 2019-nCoV that shares 91% sequence identity with SARS-CoV S2 subunit. Meanwhile, we predict the post-fusion structure of 2019-nCoV S glycoproteins through SWISS-MODEL[46] using post-fusion SARS-CoV S2 trimer as template (Supplementary Fig. 5a). We found that post-fusion S2 trimer of 2019-nCoV adopts similar architecture compared to the post-fusion SARS-CoV S2 trimer, with overall RMSD of 0.3 Å (Supplementary Fig. 5b-c). Superposition with the SARS-CoV and MHV S2 structures showed that the HR2 helix in 2019-nCoV also shift around the central HR1 helixes with similar degree of the SARS-CoV HR2 motif, and the linker

loops upstream of the HR2 motif from SARS-CoV and 2019-nCoV also bind along the stem helixes with similar conformation (Supplementary Fig. 5c).

The glycan shield of the post-fusion SARS-CoV S2 trimer

The S glycoprotein is the major antigenic determinant exposed on the viral surface. Coronavirus deploys a large number of N-linked glycans covering the S glycoprotein surface to limit the accessibility of neutralizing antibodies and evade the attack of host immune system. Our structure identified 8 out of 9 putative N-linked glcans which reveals the distribution of the Glycan shield covering the surface of the post-fusion SARS-CoV S2 trimer (Fig. 5). 4 N-linked glycosylation sites (N691, N699, N1056 and N1080) are located in the relatively stable β-sheet enriched region surrounding the C-terminal of the central stem helixes. The other 4 glycosylation sites (N1116, N1140, N1155 and N1176) are located within the HR2 motif and its upstream linker region. Conservation analysis showed that 5 N-linked glycosylation sites (N699, N1080, N1140, N1155 and N1176) are completely conserved among SARS-CoV, 2019-nCoV, MERS-CoV and MHV, whereas 3 N-linked glycosylation sites (N691, N1056 and N1116) are only conserved between SARS-CoV and 2019-nCoV (Supplementary Fig. 6). These glycans mask the accessible surface of the corresponding regions during transition from pre-fusion to post-fusion state, which may protect the S2 subunit from antibody recognition. The enrichment of glycosylate modifications in the HR2 motif and its upstream linker region indicates the importance of these regions for S2 mediated membrane fusion.

Locations of therapeutic targets within the S2 fusion machinery

Therapeutic targets in coronavirus S glycoprotein are mainly focused on the RBD region from the S1 subunit[15,16] and the HR1-HR2 six-helix bundle from the S2 subunit[17-19]. Antibodies and inhibitors binding to the RBD can prevent viral entry by inhibiting association of S glycoprotein with the receptor on the surface of the target cells. Antibodies and inhibitors against S2 subunit can inhibit virus infection by blocking the S2 meditated membrane fusion. Many peptide inhibitors of S2 subunit

are HR1 or HR2 derived peptides which prevent the fusion core formation by interacting with the corresponding HR2 or HR1 motif in the S2 fusion machinery[47,48]. The major target regions of the peptide inhibitors are residues 902-947 of the HR1 motif and residues 1149-1186 of the HR2 motif[47,48], mapped to the hydrophobic HR1-HR2 interacting groove in the post-fusion S2 trimer (Fig. 6a). Moreover, some antibodies also target the HR1-HR2 interface of SARS-CoV S2 subunit[17,21]. Besides the therapeutics toward HR1-HR2 binding groove, antibodies against other parts of the S2 subunit are also reported to block the viral entry mediated by S glycoprotein[17,18]. A previous study developed two groups of SARS-CoV mAbs targeting the region upstream of the HR2 motif, including the C-terminal of the SD3 domain and the HR2 upstream linker region[17]. Further study found that one of the antibodies targeting the epitope within the residues 1111-1130 of the HR2 upstream linker region (Fig. 6a), named mAb 1A9, interacts with D1128 from the SARS-CoV S2 subunit, and D1128A substitution enables the virus to escape the inhibition by mAb 1A9 [18]. The D1128 is located in the acidic amino acids enriched region (residues 1126-1133) of the linker loop upstream from the HR2 motif, which binds to the positive cavity in the main body as shown in our post-fusion SARS-CoV S2 structure (Fig. 3b and 6a). It is possible that mAb 1A9 could occupy D1128 and its adjacent region, which makes the linker loop unable to bind along the central helix, thus blocking the association of the HR2 motif with the HR1 helixes in the late stage of membrane fusion. Interestingly, the mAb 1A9 presents cross protection against viral entry mediated by S glycoprotein of human SARS-CoV, civet SARS-CoV and bat SL-CoVs[18]. Note that the residue D1128 and its adjacent area within the upstream linker loop are completely conserved between SARS-CoV and 2019-nCoV (Supplementary Fig. 6). Moreover, some mAbs targeting the highly conserved HR1-HR2 interacting region of SARS-CoV could also inhibit viral entry mediated by S proteins with different RBD sequences from clinical isolates[21]. According to the sequence analysis, the HR2 motif is completely conserved between SARS-CoV and 2019-nCoV, while the HR1 motif is relatively variable (Supplementary Fig. 6). These results indicate that the conserved regions of the HR2 motif and its upstream linker

loop, which are both responsible for the six-helix bundle formation as shown in our post-fusion S2 structure, are important potential targets for broad-spectrum vaccines and drugs development against SARS-like coronaviruses.

Other conserved regions within the S2 subunit may also contribute to therapeutic development. For example, the U-turn loop connecting the HR1 and CH motifs in the pre-fusion state, which refolds into the continuous HR1-CH helix during membrane fusion as shown in our structure, are accessible even in the pre-fusion S trimer when the variable RBD domain is in the standing state (Fig. 6b). This U-turn loop and its adjacent regions are completely conserved between SARS-CoV and 2019-nCoV (Supplementary Fig. 6). Moreover, other loop regions within the HR1 and CH motifs, which are also highly conserved, are totally exposed after dissociation of the S1 subunits (Fig. 6c and Supplementary Fig. 6). Small inhibitors and neutralizing antibodies against these conserved loops may also inhibit viral entry by blocking the HR1-CH stem helix formation toward a broader range of coronaviruses.

Discussion

Here we present the structure of the induced post-fusion SARS-CoV S2 machinery, which shows large scale of structural rearrangements and conformational changes compared to the structure in the pre-fusion state. Structural analysis indicates that the linker region upstream of the HR2 motif binding along the stem helix plays an important role in the subsequent HR1-HR2 six-helix bundle formation, which is critical for viral entry among class I fusion protein. The fusion mechanism revealed by the SARS-CoV S2 trimer resembles that of other class I fusion glycoproteins[49]. Based on the structural information of the pre-fusion and post-fusion SARS-CoV S glycoprotein, we proposed a model of SARS-CoV fusion process (Supplementary Fig. 7). After trypsin cleavage at the boundary of the S1 and S2 subunit, most of the S glycoproteins retained the pre-fusion trimeric architecture. Further binding with the receptor from the host membrane surface by the standing RBD accelerates the release of the S1 subunit. Then helical regions in the remained S2 trimer are refolded, which helps the fusion peptide to insert into the target membrane. The UH, BH and SD3

regions retain their structures positioned adjacent to the C-terminal of the CH helix. Next, the linker region upstream of the HR2 motif is released and extended along the axis of the stem helixes. Thus the downstream HR2 motif gets close to the HR1 helix and forms the six-helix bundle at the host membrane proximal side, which mediates fusion between the viral and host membrane. Considering the conservation of S2 subunit among coronaviruses, it is possible that all the coronaviruses mediate vial-host membrane fusion with similar strategy.

From 2003 to nowadays, the human-infecting coronaviruses SARS-CoV, MERS-CoV and 2019-nCoV have caused worldwide public health emergencies[9]. Considering the potential risk of the spread of these coronaviruses or other unknown coronaviruses in the future, as well as the frequent mutation rates in the replication of RNA viruses, developments of broad-spectrum antiviral vaccines and therapeutics are needed to defense a wide range of coronaviruses infection. Because of the important role in mediating receptor recognition and viral entry, S glycoprotein is the key determinant for efficient transmission among species and the major antigen on the surface of the coronavirus. A majority of antibodies and inhibitors are developed against the RBD region of the S1 subunit and inhibit virus infection by blocking the interaction with the host receptor. However, applications of the RBD-specific antibodies are limited by the variability of RBD protein sequences, which is consistent with different receptor usage by different coronaviruses. Moreover, mutations in RBD may not affect viral entry but impair the therapeutic efficacy of the RBD-specific treatments. For example, even though four out of five residues critical for receptor binding are replaced in the RBD domain of 2019-nCoV S glycoprotein, the 2019-nCoV still retains strong binding affinity to the SARS-CoV receptor ACE2 and infects human cells efficiently[50]. The recently published work that determined 2019-nCoV S glycoprotein in the pre-fusion state also demonstrate the binding ability of 2019-nCoV RBD with ACE2, which is even stronger than that of SARS-CoV[25]. By contrast, therapeutics developed based on the S2 subunit may benefit the treatment against a wide range of coronaviruses, since the S2 subunit is more conserved than the S1 subunit due to their critical role in mediating the subtle process of membrane

fusion (Figure S7). For example, the S1 subunit of the newly identified 2019-nCoV shares a protein sequence identity of 64% with the SARS-CoV S1 subunit, while the protein sequence identity between the S2 subunits is 91%. Meanwhile, by predicting the structure of 2019-nCoV S2 subunit, we found that the S2 subunits adopt similar structural features of the post-fusion SARS-CoV S2 machinery. Note that some SARS-CoV S2-sepcific antibodies toward the interface groove of the HR1-HR2 six-helix bundle and the linker region upstream of the HR2 motif have been demonstrated able to neutralize a broader spectrum of SARS-CoV clinical isolates or variants[18,21]. Moreover, a reported peptide inhibitor P9, which could bind with influenza HA glycoprotein and MERS-CoV S2 subunit, showed antiviral ability against a broad range of respiratory viruses, including influenza viruses, MERS-CoV and SARS-CoV[51]. Thus, we suggested that the conserved regions within the S2 subunit, including the HR2 motif, the linker region upstream of the HR2 motif and the loop regions from the HR1 and CH motifs, are important potential targets for developing broad-spectrum vaccines and therapeutics. Moreover, we described most of the glycosylate modifications of the SARS-CoV S2 fusion machinery, which may cause immune evasion by covering the protein surface and are completely conserved between SARS-CoV and 2019-nCoV. Thus our structure provides more structural information for effective antibodies and inhibitors development.

In summary, our structure of the post-fusion SARS-CoV S2 trimer, which reveals the fusion mechanism of coronavirus and the glycan shield masking the protein surface, may help antibodies development and drugs designation against SARS-CoV and other SARS-like coronaviruses, particularly the newly discovered 2019-nCoV, which is still spreading worldwide.

Acknowledgments

We thank X. Huang, B. Zhu, L. Chen and X. Li from the Center for Biological Imaging, Core Facilities for protein Science at the Institute of Biophysics (IBP), Chinese Academy of Sciences (CAS) for help with cryo-EM data collection. The project was funded by National Natural Science Foundation of China (31900871,

31570874). X.Z received scholarships from the ‘National Thousand (Young) Talents Program’ from the Office of Global Experts Recruitment in China. D.C is supported by the Youth Innovation Promotion Association of the Chinese Academy of Sciences.

Author Contributions

X.Z and D.C conceived the project; X.F performed the protein sample purification, negative staining and cyo-EM sample preparation; X.F and L.K screened the negative staining samples and processed the negative staining data. D.C and X.F screened the cryo-EM samples, collected and processed the cryo-EM data; D.C built and refined the structure model; X.Z and D.C wrote the manuscript.

References

1 Su, S. *et al.* Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends in microbiology* **24**, 490-502, doi:10.1016/j.tim.2016.03.003 (2016).

2 Lu, G., Wang, Q. & Gao, G. F. Bat-to-human: spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends in microbiology* **23**, 468-478, doi:10.1016/j.tim.2015.06.003 (2015).

3 Weinstein, R. A. Planning for epidemics--the lessons of SARS. *The New England journal of medicine* **350**, 2332-2334, doi:10.1056/NEJMp048082 (2004).

4 Azhar, E. I., Hui, D. S. C., Memish, Z. A., Drosten, C. & Zumla, A. The Middle East Respiratory Syndrome (MERS). *Infectious disease clinics of North America* **33**, 891-905, doi:10.1016/j.idc.2019.08.001 (2019).

5 Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. & Fouchier, R. A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *The New England journal of medicine* **367**, 1814-1820, doi:10.1056/NEJMoa1211721 (2012).

6 Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *The New England journal of medicine* **382**, 727-733, doi:10.1056/NEJMoa2001017 (2020).

7 Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. doi:10.1038/s41586-020-2008-3 (2020).

8 She, J. *et al.* 2019 novel coronavirus of pneumonia in Wuhan, China: emerging attack and management strategies. *Clinical and translational medicine* **9**, 19, doi:10.1186/s40169-020-00271-z (2020).

9 Chen, Y., Liu, Q. & Guo, D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. **92**, 418-423, doi:10.1002/jmv.25681 (2020).

10 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, doi:10.1038/s41586-020-2012-7 (2020).

11 Bosch, B. J., van der Zee, R., de Haan, C. A. & Rottier, P. J. The coronavirus spike protein is a

class I virus fusion protein: structural and functional characterization of the fusion core complex. *Journal of virology* **77**, 8801-8811, doi:10.1128/jvi.77.16.8801-8811.2003 (2003).

12 Beniac, D. R., deVarennes, S. L., Andonov, A., He, R. & Booth, T. F. Conformational reorganization of the SARS coronavirus spike following receptor binding: implications for membrane fusion. *PloS one* **2**, e1082, doi:10.1371/journal.pone.0001082 (2007).

13 Belouzard, S., Chu, V. C. & Whittaker, G. R. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5871-5876, doi:10.1073/pnas.0809524106 (2009).

14 Du, L. *et al.* The spike protein of SARS-CoV--a target for vaccine and therapeutic development. *Nature reviews. Microbiology* **7**, 226-236, doi:10.1038/nrmicro2090 (2009).

15 Walls, A. C. *et al.* Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **176**, 1026-1039.e1015, doi:10.1016/j.cell.2018.12.028 (2019).

16 Sui, J. *et al.* Potent neutralization of severe acute respiratory syndrome (SARS) coronavirus by a human mAb to S1 protein that blocks receptor association. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 2536-2541, doi:10.1073/pnas.0307140101 (2004).

17 Lip, K. M. *et al.* Monoclonal antibodies targeting the HR2 domain and the region immediately upstream of the HR2 of the S protein neutralize in vitro infection of severe acute respiratory syndrome coronavirus. *Journal of virology* **80**, 941-950, doi:10.1128/jvi.80.2.941-950.2006 (2006).

18 Ng, O. W. *et al.* Substitution at aspartic acid 1128 in the SARS coronavirus spike glycoprotein mediates escape from a S2 domain-targeting neutralizing monoclonal antibody. *PloS one* **9**, e102415, doi:10.1371/journal.pone.0102415 (2014).

19 Chu, L. H. *et al.* Fusion core structure of the severe acute respiratory syndrome coronavirus (SARS-CoV): in search of potent SARS-CoV entry inhibitors. *Journal of cellular biochemistry* **104**, 2335-2347, doi:10.1002/jcb.21790 (2008).

20 Coughlin, M. M. & Prabhakar, B. S. Neutralizing human monoclonal antibodies to severe acute respiratory syndrome coronavirus: target, mechanism of action, and therapeutic potential. *Reviews in medical virology* **22**, 2-17, doi:10.1002/rmv.706 (2012).

21 Elshabrawy, H. A., Coughlin, M. M., Baker, S. C. & Prabhakar, B. S. Human monoclonal antibodies against highly conserved HR1 and HR2 domains of the SARS-CoV spike protein are more broadly neutralizing. *PloS one* **7**, e50366, doi:10.1371/journal.pone.0050366 (2012).

22 Yuan, Y. *et al.* Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nature communications* **8**, 15092, doi:10.1038/ncomms15092 (2017).

23 Gui, M. *et al.* Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell research* **27**, 119-129, doi:10.1038/cr.2016.152 (2017).

24 Pallesen, J. *et al.* Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. **114**, E7348-e7357, doi:10.1073/pnas.1707304114 (2017).

25 Wrapp, D. & Wang, N. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. doi:10.1126/science.abb2507 (2020).

26 Song, W., Gui, M., Wang, X. & Xiang, Y. Cryo-EM structure of the SARS coronavirus spike

glycoprotein in complex with its host cell receptor ACE2. **14**, e1007236, doi:10.1371/journal.ppat.1007236 (2018).

27 Kirchdoerfer, R. N., Wang, N. & Pallesen, J. Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. **8**, 15701, doi:10.1038/s41598-018-34171-7 (2018).

28 Xu, Y. *et al.* Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. *The Journal of biological chemistry* **279**, 49414-49419, doi:10.1074/jbc.M408782200 (2004).

29 Supekar, V. M. *et al.* Structure of a proteolytically resistant core from the severe acute respiratory syndrome coronavirus S2 fusion protein. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17958-17963, doi:10.1073/pnas.0406128102 (2004).

30 Gao, J. *et al.* Structure of the fusion core and inhibition of fusion by a heptad repeat peptide derived from the S protein of Middle East respiratory syndrome coronavirus. *Journal of virology* **87**, 13134-13140, doi:10.1128/jvi.02433-13 (2013).

31 Xu, Y. *et al.* Structural basis for coronavirus-mediated membrane fusion. Crystal structure of mouse hepatitis virus spike protein fusion core. *The Journal of biological chemistry* **279**, 30514-30522, doi:10.1074/jbc.M403760200 (2004).

32 Walls, A. C. *et al.* Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* **531**, 114-117, doi:10.1038/nature16988 (2016).

33 Walls, A. C. *et al.* Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 11157-11162, doi:10.1073/pnas.1708727114 (2017).

34 Wu, C., Huang, X., Cheng, J., Zhu, D. & Zhang, X. High-quality, high-throughput cryo-electron microscopy data collection via beam tilt and astigmatism-free beam-image shift. *Journal of structural biology* **208**, doi:10.1016/j.jsb.2019.09.013 (2019).

35 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).

36 Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of structural biology* **192**, 216-221, doi:10.1016/j.jsb.2015.08.008 (2015).

37 Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of structural biology* **180**, 519-530, doi:10.1016/j.jsb.2012.09.006 (2012).

38 Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nature methods* **11**, 63-65, doi:10.1038/nmeth.2727 (2014).

39 Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. **10**, 845-858, doi:10.1038/nprot.2015.053 (2015).

40 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).

41 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501, doi:10.1107/s0907444910007493 (2010).

42 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213-221, doi:10.1107/s0907444909052925 (2010).

43 Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37-43, doi:10.1038/371037a0 (1994).

44 Yin, H. S., Wen, X., Paterson, R. G., Lamb, R. A. & Jardetzky, T. S. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. *Nature* **439**, 38-44, doi:10.1038/nature04322 (2006).

45 Walls, A. C. *et al.* Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nature structural & molecular biology* **23**, 899-905, doi:10.1038/nsmb.3293 (2016).

46 Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research* **46**, W296-w303, doi:10.1093/nar/gky427 (2018).

47 Liu, I. J. *et al.* Identification of a minimal peptide derived from heptad repeat (HR) 2 of spike protein of SARS-CoV and combination of HR1-derived peptides as fusion inhibitors. *Antiviral research* **81**, 82-87, doi:10.1016/j.antiviral.2008.10.001 (2009).

48 Zhu, J. *et al.* Following the rule: formation of the 6-helix bundle of the fusion core from severe acute respiratory syndrome coronavirus spike protein and identification of potent peptide inhibitors. *Biochemical and biophysical research communications* **319**, 283-288, doi:10.1016/j.bbrc.2004.04.141 (2004).

49 Harrison, S. C. Viral membrane fusion. *Virology* **479-480**, 498-507, doi:10.1016/j.virol.2015.03.043 (2015).

50 Xu, X. *et al.* Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Science China. Life sciences*, doi:10.1007/s11427-020-1637-5 (2020).

51 Zhao, H. *et al.* A novel peptide with potent and broad-spectrum antiviral activities against multiple respiratory viruses. *Scientific reports* **6**, 22008, doi:10.1038/srep22008 (2016).

Figure legend

Figure 1 Overall structure of SARS-CoV S2 fusion machinery in post-fusion state.

(**a**) Schematic diagram of the SARS-CoV S2 subunit. The uncolored region represents the undetermined region in our structure. L, linker region; UH, the upstream helixes; FP, fusion peptide; CR, connecting region; HR, heptad repeat; CH, central helix; BH, β-hairpin; SD3, subdomain 3. (**b-e**) Cartoon representation of the SARS-CoV S2 trimer in post-fusion state colored as shown in a.

Figure 2 Structural and conformational changes of SARS-CoV S glycoprotein from pre-fusion to post-fusion state.

(**a**) Topology and cartoon representation of SARS-CoV S2 subunit in pre-fusion state colored as in Fig. 1a. The S1 subunit in the pre-fusion state is colored white. (**b**) Topology and cartoon representation of SARS-CoV S2 subunit in post-fusion state colored as in Fig. 1a. Unobserved regions in the structures are colored in light gray or presented as dashed lines in the topology diagram of a and b. (**c**) Structural transition of the HR1 and CH motifs during membrane fusion. (**d**) Structural changes of the UH, BH and SD3 during membrane fusion.

Figure 3 The linker region upstream of HR2 motif important for membrane fusion.

(**a**) binding of the linker region upstream of HR2 motif along the central helix colored as in Fig. 1a. (**b**) The electrostatic surface representation of the linker region upstream of HR2 motif binding area. Red represents negative charge, and blue represents positive charge. (**c**) Overall structural changes of the linker region upstream of the HR2 motif during membrane fusion. The linker regions in pre-fusion and post-fusion states are shown as cartoons colored in violet and green, respectively. The SD3 domains in pre-fusion and post-fusion states are shown as cartoons colored in pale cyan and orange, respectively. (**d**) Structural changes of the linker region upstream of the HR2 motif during membrane fusion represented in single chain mode. The cartoons are colored as in c. (**e**) Structural details of the linker region and the SD3 domain from the pre-fusion SARS-CoV colored as in c. (**f**) Structural details of the linker region and the SD3 domain from the post-fusion SARS-CoV colored as in c. The disulfide bond is shown as spheres.

Figure 4 Comparison of SARS-CoV and MHV S2 structures in post-fusion state.

(**a-b**) Cartoon representation of the comparison of post-fusion S2 overall structures from SARS-CoV (cyan) and MHV (wheat). (**c**) Superposition of SARS-CoV (colored as Fig. 1a) and MHV (wheat) post-fusion S2 subunits in single chain mode. (**d-e**)

Structural comparison of the linker region upstream of HR2 motif from SARS-CoV (green) and MHV (wheat) shown as cartoon and electrostatic surface. Red represents negative charge, and blue represents positive charge.

Figure 5 Glycan shield of SARS-CoV S2 fusion machinery in post-fusion state.

(**a**) Schematic diagram of the N-linked glycosylate modification sites of SARS-CoV S2 subunit. The visible N-linked glycosylation sites are shown in solid lines and the invisible N-linked glycosylation site is shown in dashed line. (**b-d**) The distribution of the glycan shield within the S2 subunit shown in side view (b), top view (c) and bottom view (d). The highly conserved glycosylation sites among coronavirus are shown as light blue spheres. The completely conserved glycosylation sites between SARS-CoV and 2019-nCoV are shown as yellow spheres.

Figure 6 Potential therapeutic targets within SARS-CoV S2 subunit.

(**a**) Cartoon representation of the linker loop region upstream of the HR2 motif (blue loop), the HR1 motif (blue helix) and the HR2 motif (red). The residue D1128 is shown as spheres colored in red. (**b**) The U-turn loop connecting the HR1 and CH motif in the pre-fusion SARS-CoV S trimer. The pre-fusion SARS-CoV structure is shown as cartoons colored in white. The U-turn loop is presented as spheres colored in red. (**c**) The U-turn loop and other loop regions within the HR1 and CH motifs in the pre-fusion S2 subunit. The U-turn loop is shown as spheres colored in red and the other loop regions is shown as spheres colored in blue.

Table 1. Cryo-EM data collection, refinement and validation statistics

	Post-fusion SARS-CoV S2 trimer
Data collection and processing	
Magnification	130,000
Voltage (kV)	200
Electron exposure (e^{-}/Å^2)	50
Defocus range (μm)	1.8-2.2
Pixel size (Å)	1
Symmetry imposed	C3
Initial particle images (no.)	343,042
Final particle images (no.)	40,600
Map resolution (Å)	3.9
Refinement	
Model resolution (Å)	3.9
Map sharpening *B* factor (Å^2)	-165
R.m.s. deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.32
Validation	
Clashscore	9.78
Poor rotamers (%)	0.66
Ramachandran plot	
Favored (%)	88.92
Allowed (%)	11.08
Disallowed (%)	0

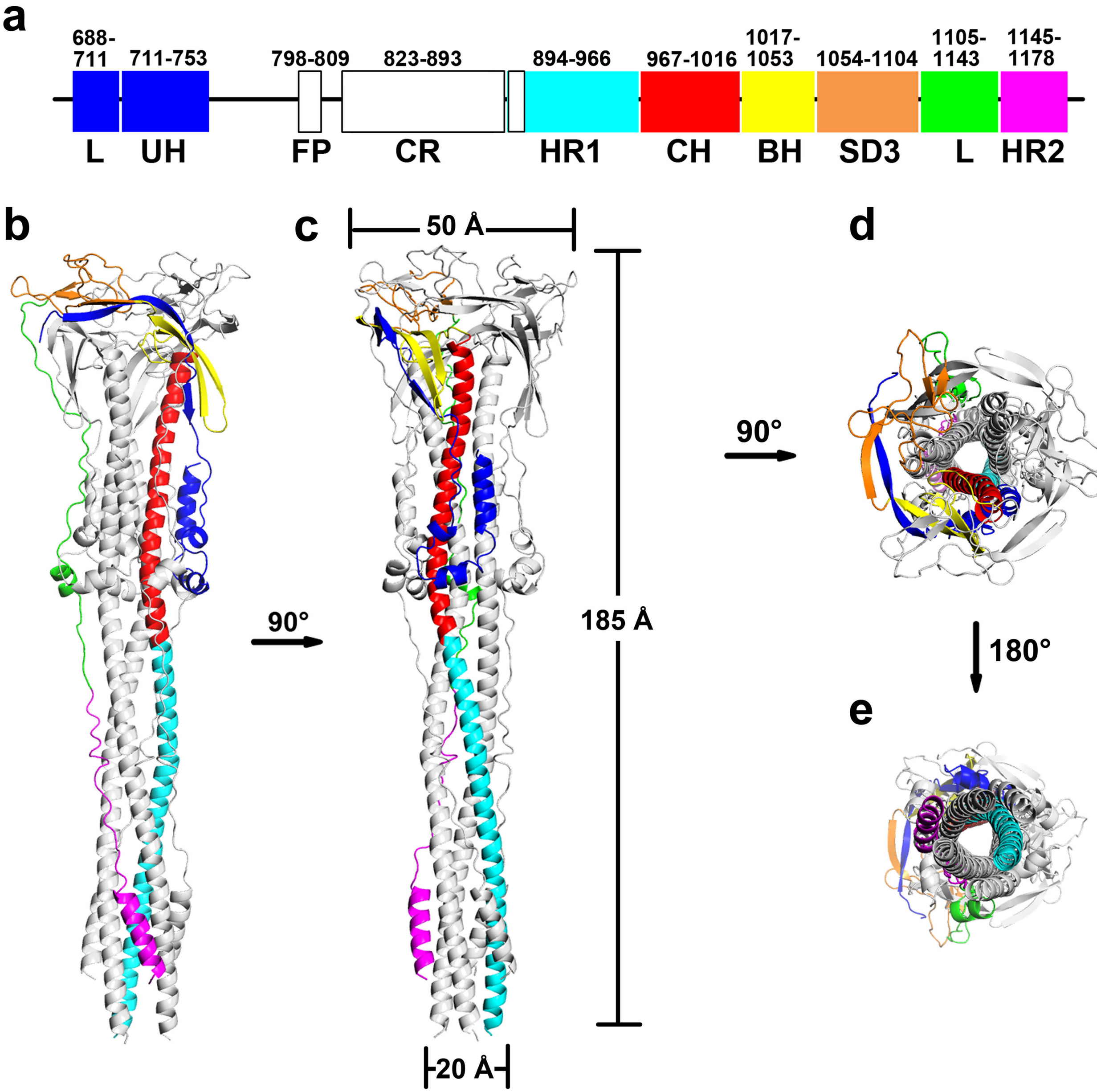

Figure 1 Overall structure of SARS-CoV S2 fusion machinery in post-fusion state.

(**a**) Schematic diagram of the SARS-CoV S2 subunit. The uncolored region represents the undetermined region in our structure. L, linker region; UH, the upstream helixes; FP, fusion peptide; CR, connecting region; HR, heptad repeat; CH, central helix; BH, β-hairpin; SD3, subdomain 3. (**b-e**) Cartoon representation of the SARS-CoV S2 trimer in post-fusion state colored as shown in a.

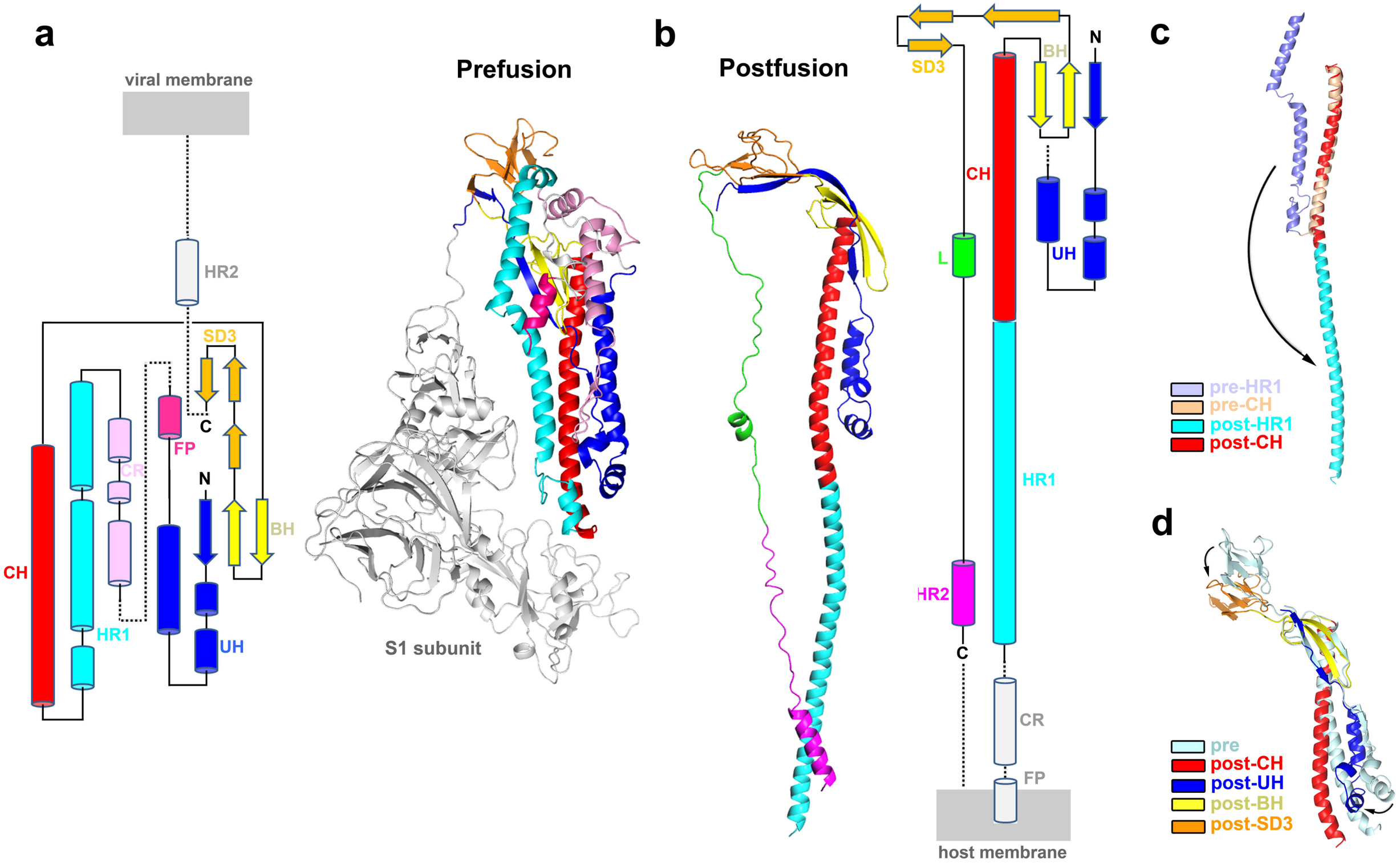

Figure 2 Structural and conformational changes of SARS-CoV S glycoprotein from pre-fusion to post-fusion state.

(**a**) Topology and cartoon representation of SARS-CoV S2 subunit in pre-fusion state colored as in Fig. 1a. The S1 subunit in the pre-fusion state is colored white. (**b**) Topology and cartoon representation of SARS-CoV S2 subunit in post-fusion state colored as in Fig. 1a. Unobserved regions in the structures are colored in light gray or presented as dashed lines in the topology diagram of a and b. (**c**) Structural transition of the HR1 and CH motifs during membrane fusion. (**d**) Structural changes of the UH, BH and SD3 during membrane fusion.

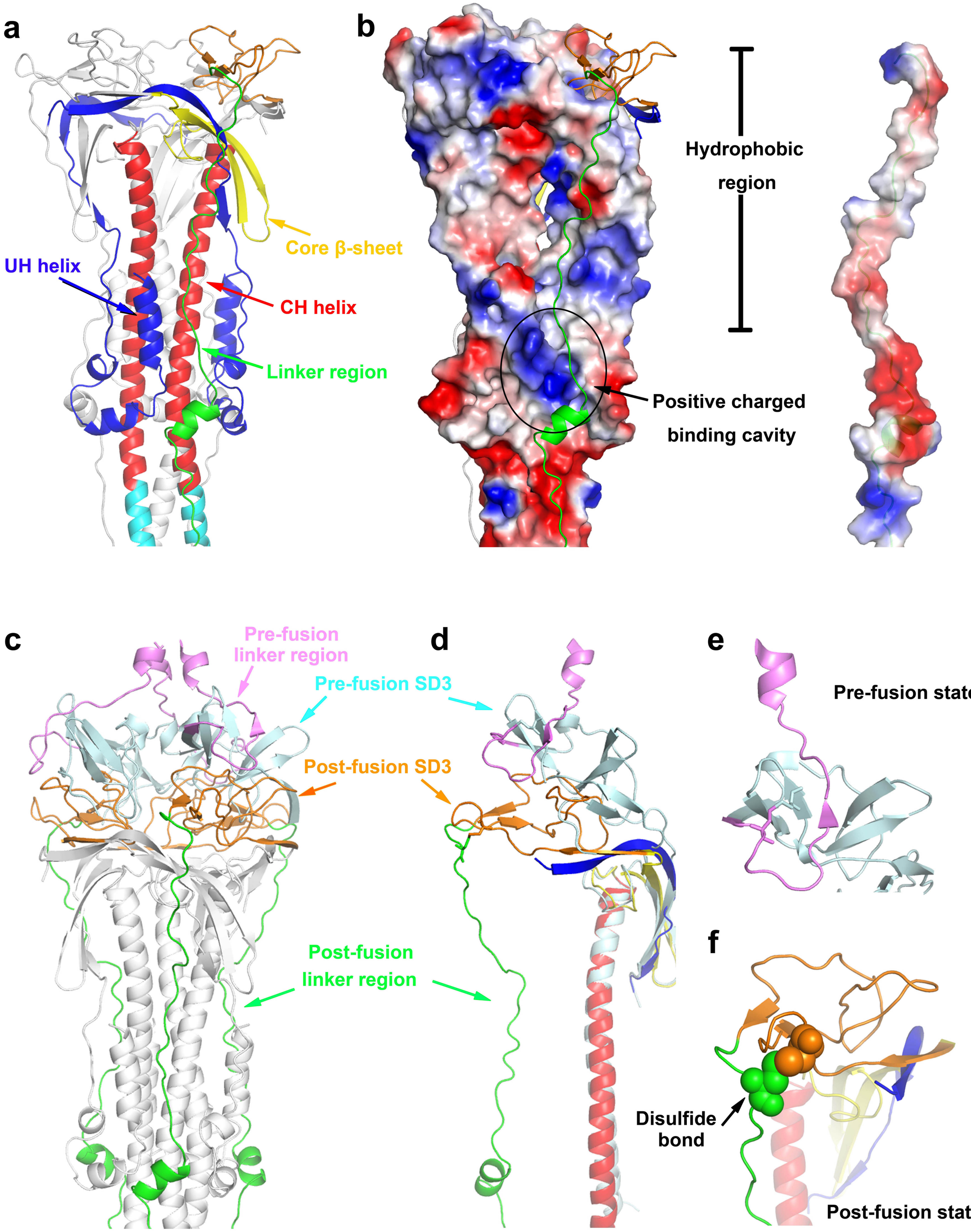

Figure 3 The linker region upstream of HR2 motif important for membrane fusion.

(**a**) binding of the linker region upstream of HR2 motif along the central helix colored as in Fig. 1a. (**b**) The electrostatic surface representation of the linker region upstream of HR2 motif binding area. Red represents negative charge, and blue represents positive charge. (**c**) Overall structural changes of the linker region upstream of the HR2 motif during membrane fusion. The linker regions in pre-fusion and post-fusion states are shown as cartoons colored in violet and green, respectively. The SD3 domains in pre-fusion and post-fusion states are shown as cartoons colored in pale cyan and orange, respectively. (**d**) Structural changes of the linker region upstream of the HR2 motif during membrane fusion represented in single chain mode. The cartoons are colored as in c. (**e**) Structural details of the linker region and the SD3 domain from the pre-fusion SARS-CoV colored as in c. (**f**) Structural details of the linker region and the SD3 domain from the post-fusion SARS-CoV colored as in c. The disulfide bond is shown as spheres.

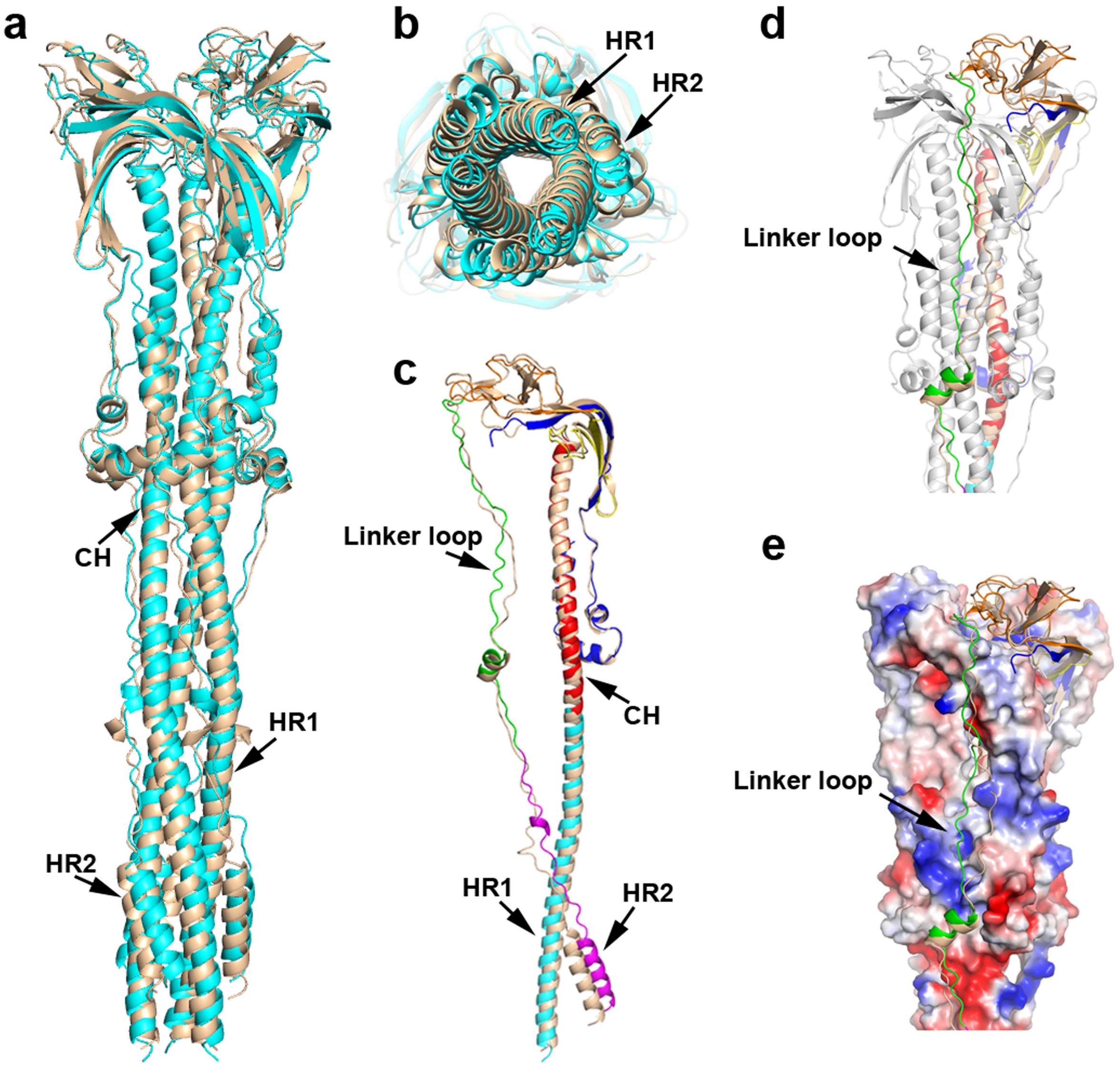

Figure 4 Comparison of SARS-CoV and MHV S2 structures in post-fusion state.

(**a-b**) Cartoon representation of the comparison of post-fusion S2 overall structures from SARS-CoV (cyan) and MHV (wheat). (**c**) Superposition of SARS-CoV (colored as Fig. 1a) and MHV (wheat) post-fusion S2 subunits in single chain mode. (**d-e**) Structural comparison of the linker region upstream of HR2 motif from SARS-CoV (green) and MHV (wheat) shown as cartoon and electrostatic surface. Red represents negative charge, and blue represents positive charge.

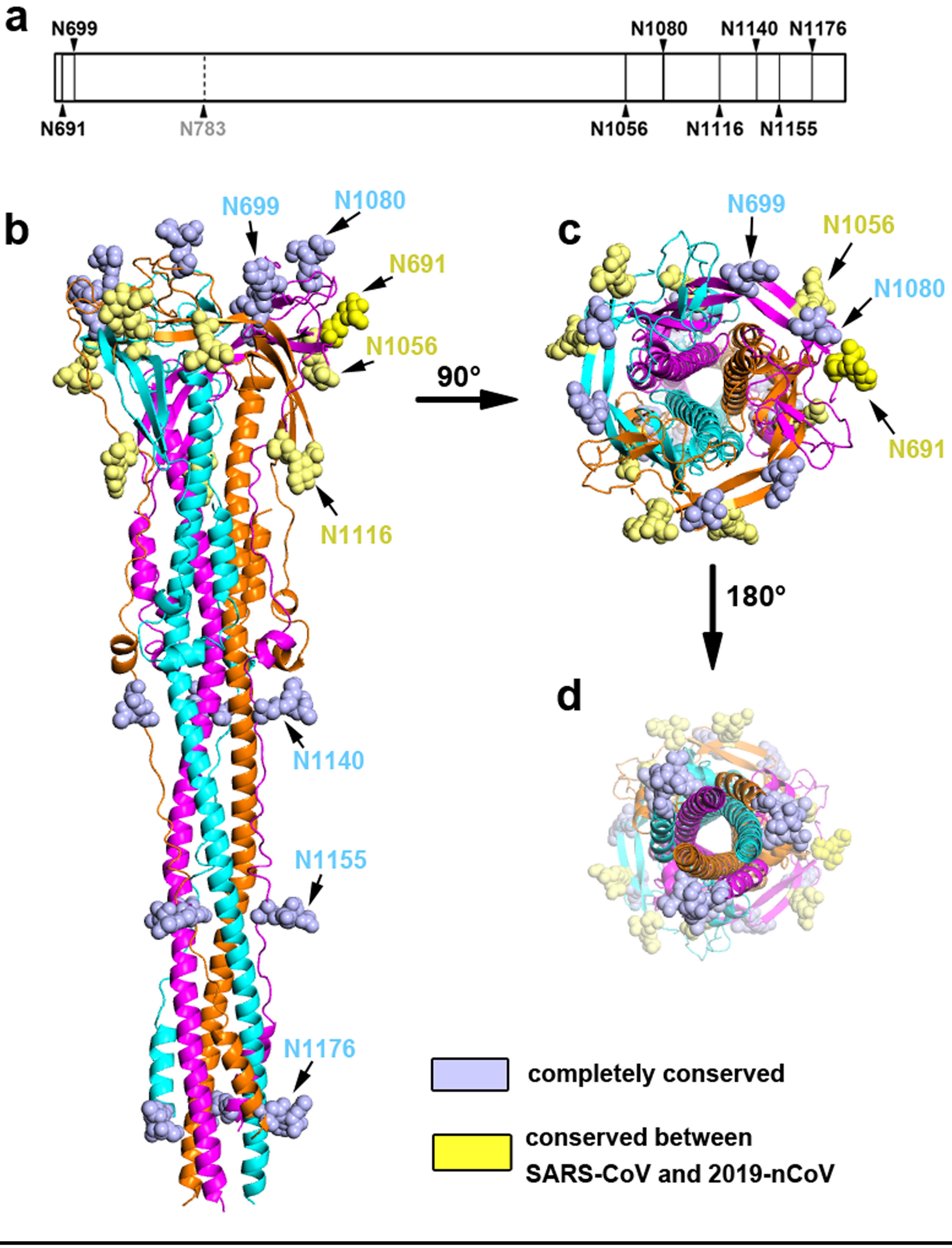

Figure 5 Glycan shield of SARS-CoV S2 fusion machinery in post-fusion state.

(**a**) Schematic diagram of the N-linked glycosylate modification sites of SARS-CoV S2 subunit. The visible N-linked glycosylation sites are shown in solid lines and the invisible N-linked glycosylation site is shown in dashed line. (**b-d**) The distribution of the glycan shield within the S2 subunit shown in side view (b), top view (c) and bottom view (d). The highly conserved glycosylation sites among coronavirus are shown as light blue spheres. The completely conserved glycosylation sites between SARS-CoV and 2019-nCoV are shown as yellow spheres.

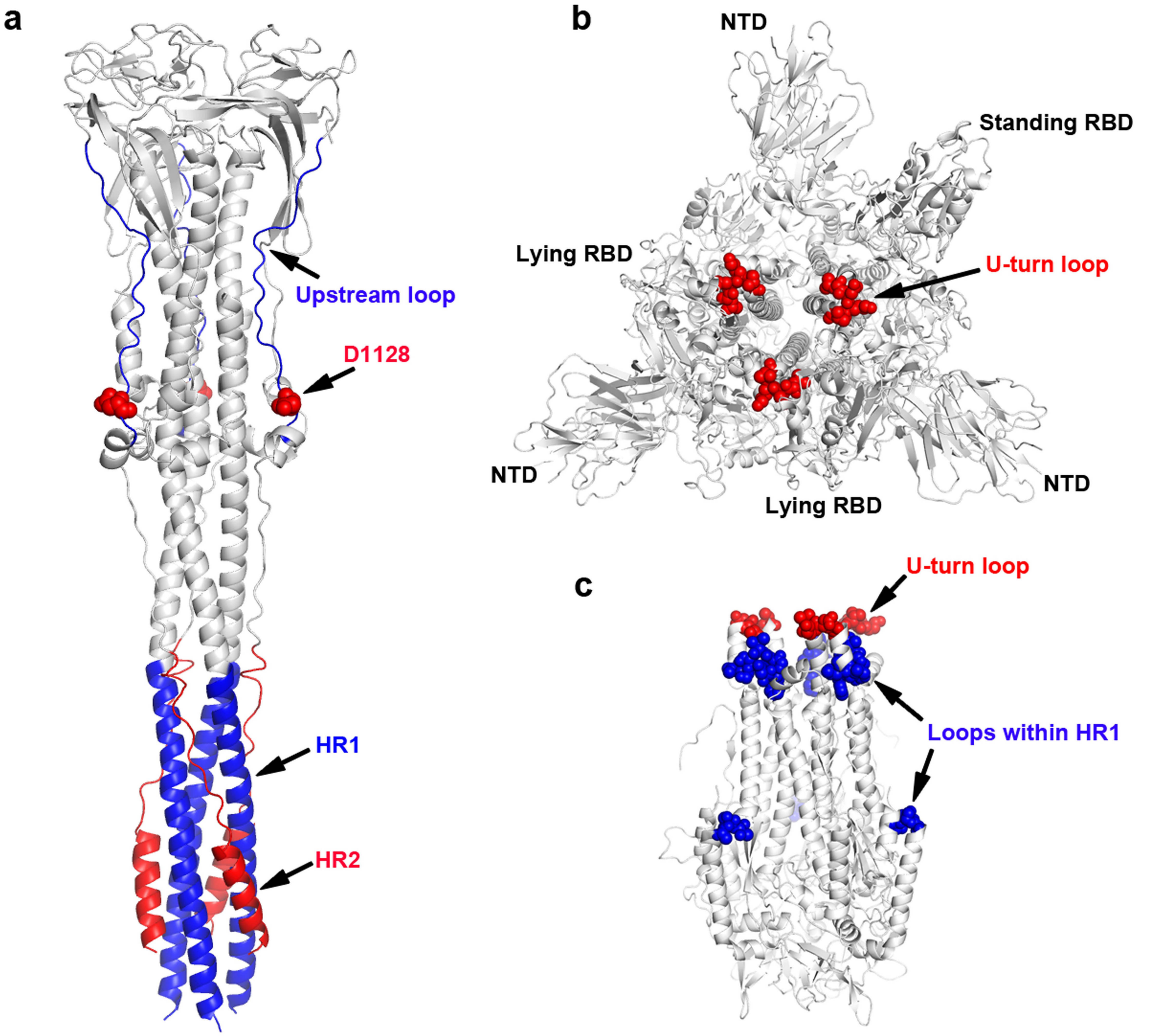

Figure 6 Potential therapeutic targets within SARS-CoV S2 subunit.

(**a**) Cartoon representation of the linker loop region upstream of the HR2 motif (blue loop), the HR1 motif (blue helix) and the HR2 motif (red). The residue D1128 is shown as spheres colored in red. (**b**) The U-turn loop connecting the HR1 and CH motif in the pre-fusion SARS-CoV S trimer. The pre-fusion SARS-CoV structure is shown as cartoons colored in white. The U-turn loop is presented as spheres colored in red. (**c**) The U-turn loop and other loop regions within the HR1 and CH motifs in the pre-fusion S2 subunit. The U-turn loop is shown as spheres colored in red and the other loop regions is shown as spheres colored in blue.